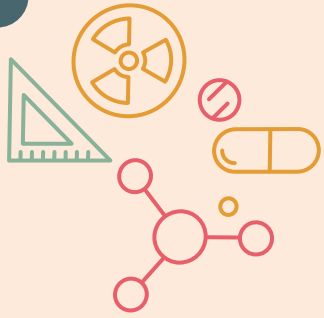




Universidad de Jaén

Tema 5: Análisis Estadístico

Grado en Educación Infantil
Asignatura: Metodología de la Investigación en
Educación Infantil
Profesor: Samuel Parra León



Índice de contenidos



01

Análisis descriptivo

Capítulo 8 & 13; Patoja, A. (2009). Manual básico para la realización de tesinas, tesis y trabajos de investigación. EOS Gabinete de Orientación Psicológica

02

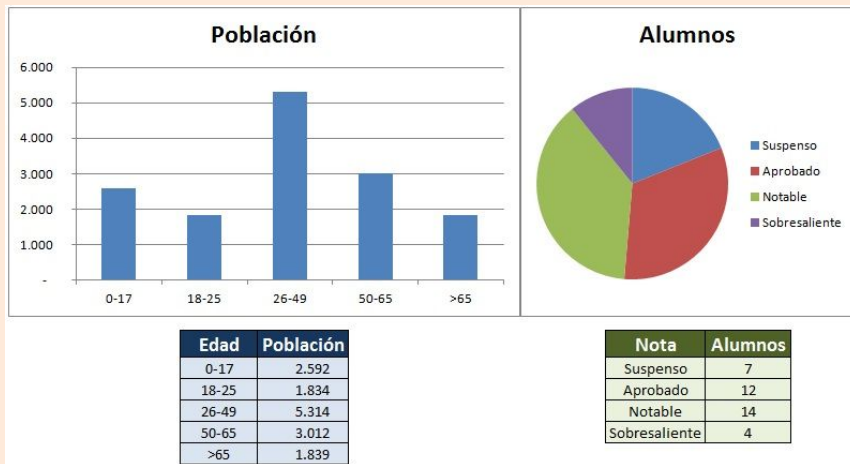
Contraste de Hipótesis.



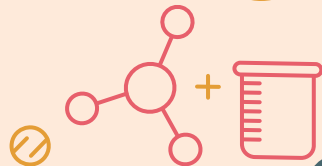
01.-Estadística Descriptiva. Introducción

¿Qué es la estadística descriptiva?

La estadística descriptiva es la rama de la estadística que recolecta, analiza y caracteriza un conjunto de datos con el objetivo de **describir** las características y comportamientos de este **conjunto mediante medidas de resumen, tablas o gráficos.**



Carácter exploratorio



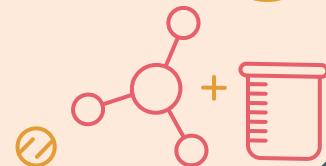
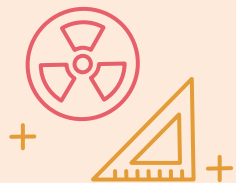
✧ 01.-Estadística Descriptiva. Introducción



¿Variables estadísticas?

Una variable estadística es el conjunto de valores que puede tomar cierta característica de la población sobre la que se realiza el estudio estadístico y sobre la que es posible su medición. Según su medición existen dos tipos de variables.

- **Cualitativa** (o categórica): son las variables que pueden tomar como valores cualidades o categorías. Ejemplos:
 - ✓ Sexo (hombre, mujer)
 - ✓ Salud (buena, regular, mala)
- **Cuantitativas** (o numérica): variables que toman valores numéricos. Ejemplos:
 - ✓ Número de casas (1, 2,...). Discreta.
 - ✓ Edad (12,5; 24,3; 35;...). Continua



* 01.-Estadística Descriptiva.

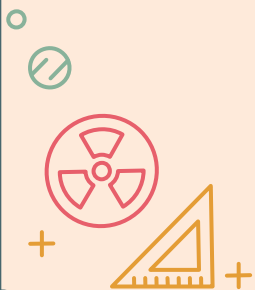
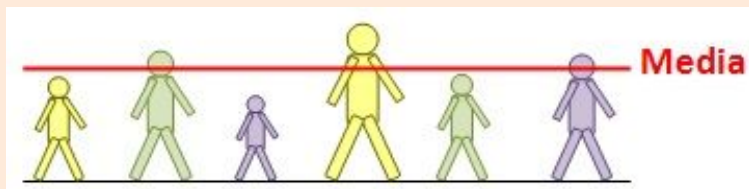
Medidas de posición central



Las **medidas de tendencia central** (o de centralización) son medidas que tienden a localizar en qué punto se encuentra la **parte central de un conjunto ordenado de datos** de una variable cuantitativa.

Media: (también llamada promedio o media aritmética) de un conjunto de datos (X_1, X_2, \dots, X_N) al valor característico de una serie de datos resultado de la suma de todas las observaciones dividido por el número total de datos

$$\text{Media}(X) = \bar{x} = \frac{X_1 + X_2 + \dots + X_N}{N}$$



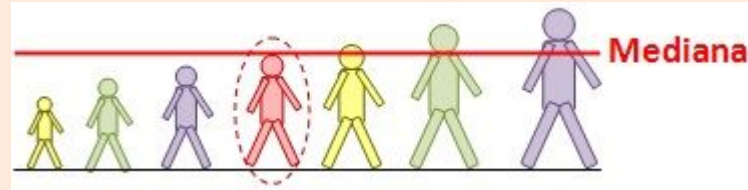


01.-Estadística Descriptiva.



Medidas de posición central

Mediana: es el elemento de un conjunto de datos ordenados (X_1, X_2, \dots, X_N) que deja a izquierda y derecha la mitad de valores.



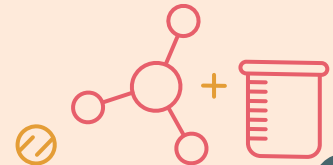
Si el conjunto de datos no está ordenado, la mediana es el valor del conjunto tal que el 50% de los elementos son menores o iguales y el otro 50% mayores o iguales.



+



+



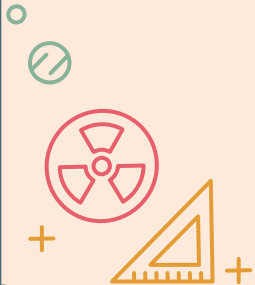
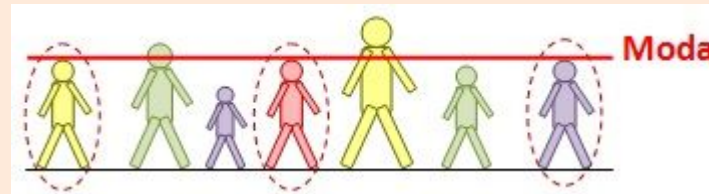


01.-Estadística Descriptiva.

Medidas de posición central



Moda: es el valor más repetido del conjunto de datos, es decir, el valor cuya frecuencia relativa es mayor. En un conjunto puede haber más de una moda.



✧ 01.-Estadística Descriptiva.



Medidas de posición NO central

Las **medidas de posición no central** (o medidas de tendencia no central) permiten conocer puntos característicos de una serie de valores, que no necesariamente tienen que ser centrales. La intención de estas medidas es **dividir** el conjunto de observaciones en grupos con el **mismo número de valores**.

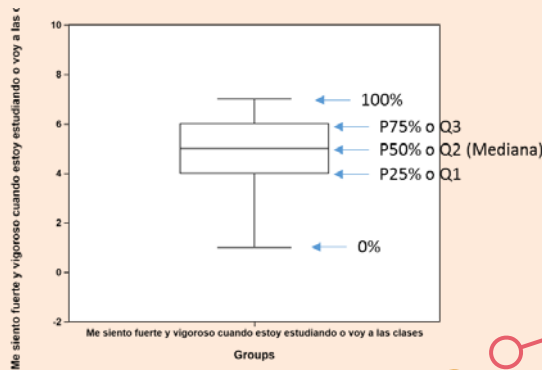
Cuartiles: son los tres valores que dividen una serie de datos ordenada en cuatro porciones iguales. El primer cuartil (Q_1) deja a la izquierda el 25% de los datos. El segundo (Q_2) deja a izquierda y derecha el 50% y coincide con la mediana. El tercero (Q_3) deja a la derecha el 25% de valores. Los tres cuartiles son:

$$\text{Cuartil}_1 = Q_1 = X_{((N+1)/4)}$$

$$\text{Cuartil}_2 = Q_2 = \text{Mediana}(X) = X_{((N+1)/2)}$$

$$\text{Cuartil}_3 = Q_3 = X_{(3(N+1)/4)}$$

siendo (X_1, X_2, \dots, X_N) la serie de datos ordenada



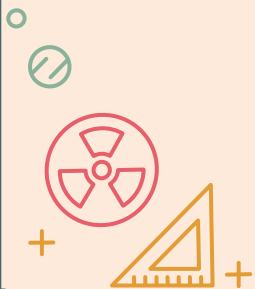
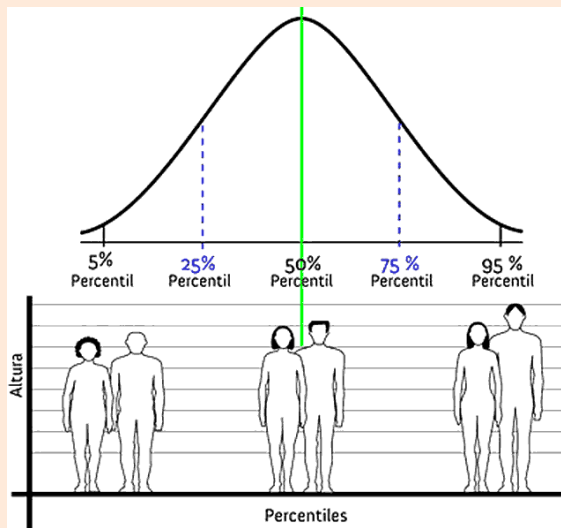
✧ 01.-Estadística Descriptiva.



Medidas de posición NO central

Percentiles: es una medida de posición no central. Los percentiles P_i son los 99 puntos que dividen una serie de datos ordenada en **100 partes iguales**, es decir, que contienen el mismo número de elementos cada una. El **percentil 50 es la mediana**.

Los percentiles están pensados para conjuntos de elementos de más de cien elementos.



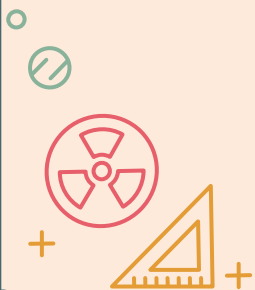
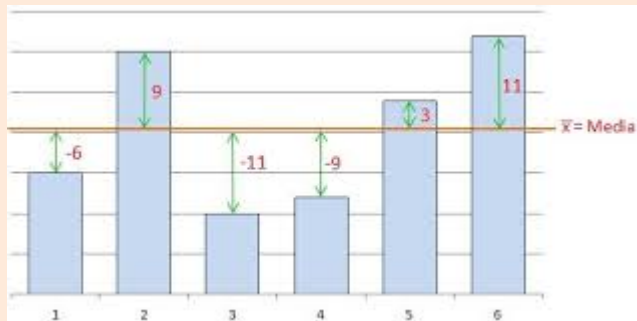
01.-Estadística Descriptiva.

Medidas de dispersión

Las medidas de dispersión o medidas de variabilidad muestran la **variabilidad de un conjunto de datos**, indicando la mayor o menor concentración de datos respecto a las medias de centralización.

Rango: (R) o recorrido estadístico es la diferencia entre el valor máximo y el mínimo de un conjunto de elementos.

Varianza: (S^2) mide la **dispersión** de los datos de una muestra **respecto a la media**, calculando la media de los cuadrados de las distancias de todos los datos.



✧ 01.-Estadística Descriptiva.

Medidas de dispersión

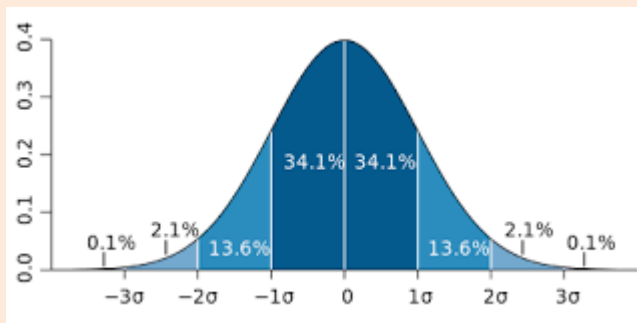


Las medidas de dispersión o medidas de variabilidad **muestran la variabilidad de un conjunto de datos**, indicando la mayor o menor concentración de datos respecto a las medias de centralización.

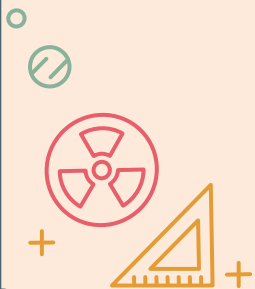
Rango: (R) o recorrido estadístico es la diferencia entre el valor máximo y el mínimo de un conjunto de elementos.

Varianza: (S^2) mide la **dispersión** de los datos de una muestra **respecto a la media**, calculando la media de los cuadrados de las distancias de todos los datos.

Desviación típica: es la medida de dispersión (S) asociada a la media. Mide el promedio de las desviaciones de los datos respecto a la media en las mismas unidades de los datos.



El cuadrado de la desviación típica es la varianza.



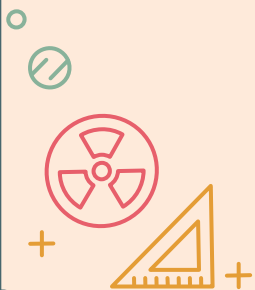
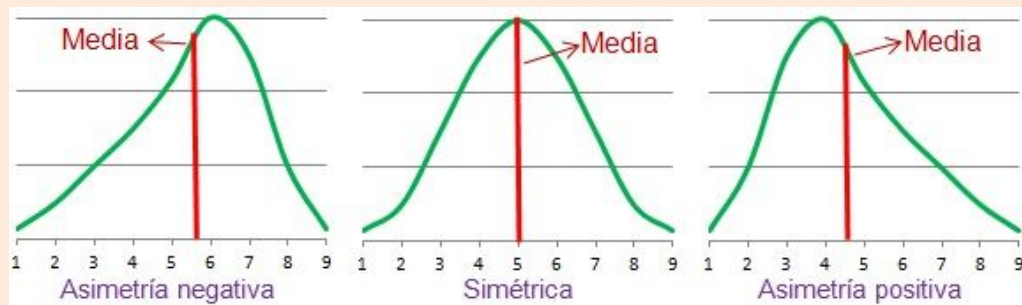
01.-Estadística Descriptiva.

Medidas de dispersión

Asimetría y curtosis, informan sobre la forma de la distribución de una variable. Estas medidas permiten saber las características de su asimetría y homogeneidad sin necesidad de representarlos gráficamente.

Asimetría: es la medida que indica la simetría de la **distribución** de una variable **respecto a la media aritmética**. Los coeficientes de asimetría indican si hay el mismo número de elementos a izquierda y derecha de la media.

- **Asimetría negativa:** la cola de la distribución se alarga para valores inferiores a la media.
- **Simétrica:** hay el mismo número de elementos a izquierda y derecha de la media. En este caso, coinciden la media, la mediana y la moda. La distribución se adapta a la forma de la campana de Gauss, o distribución normal.
- **Asimetría positiva:** la cola de la distribución se alarga para valores superiores a la media



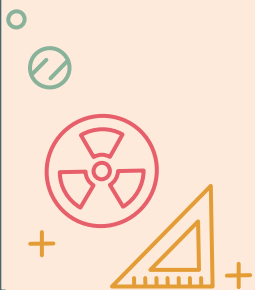
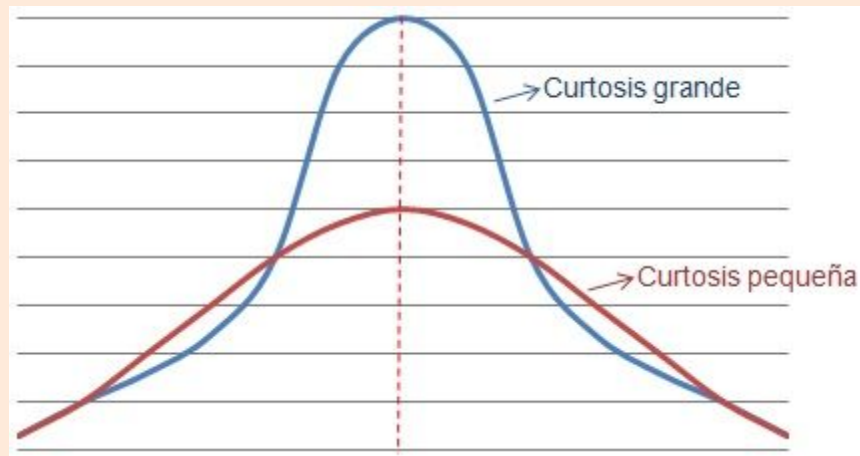
* 01.-Estadística Descriptiva.

Medidas de dispersión



Curtosis: (o apuntamiento) es una medida de forma que mide cuán escarpada o achatada está una curva o distribución.

Este coeficiente indica la **cantidad de datos que hay cercanos a la media**, de manera que a mayor grado de curtosis, más escarpada (o apuntada) será la forma de la curva.



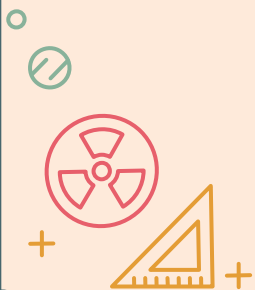
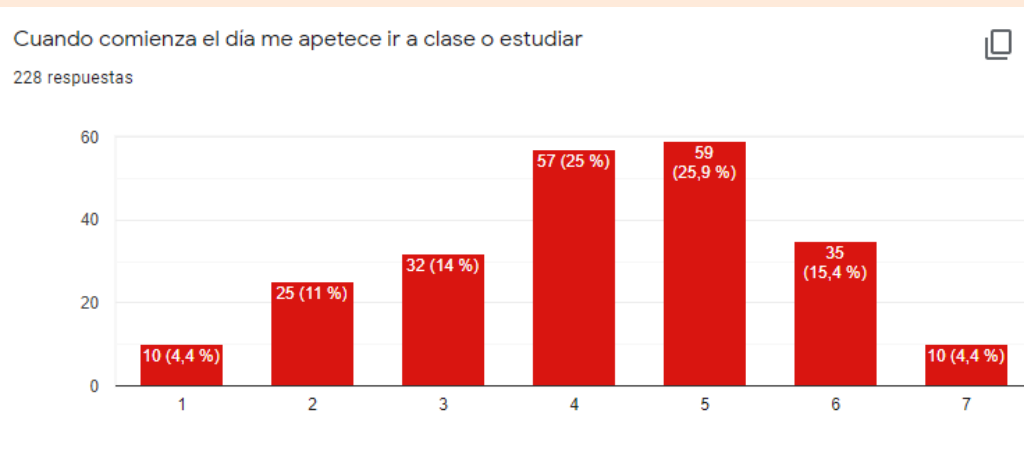
✧ 01.-Estadística Descriptiva.

Frecuencias



La **frecuencia** es una medida que sirve para comparar la aparición de un elemento X_i en un conjunto de elementos (X_1, X_2, \dots, X_N) . Mediante tablas de distribuciones de frecuencia se puede presentar organizadamente el recuento de datos.

Las frecuencias de cada elemento se pueden expresar tanto **absolutas** (número total de apariciones) como **relativas** (proporción de apariciones).





01.-Estadística Descriptiva. Gráficos



¿Un **gráfico** (o gráfica) es el recurso de representar los datos numéricos por medio de líneas, diagramas, dibujos, etc. La representación gráfica es un importante suplemento al análisis y estudio estadístico.

Los gráficos llaman la atención del lector y hacen que de un vistazo éste tenga una mayor comprensión de los datos. Un buen gráfico puede captar al lector para que a continuación lea todo el estudio. Si un estudio se compone únicamente de texto y tablas, posiblemente no todos los lectores lean el estudio.

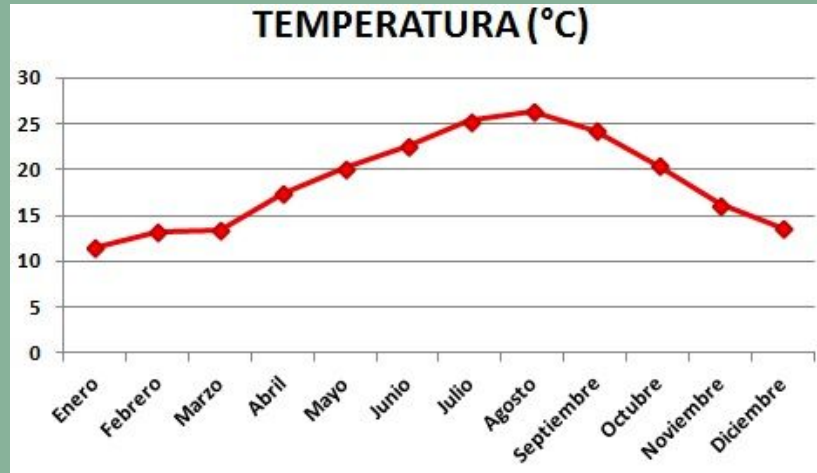




01.-Estadística Descriptiva. Gráficos



Gráfico lineal: se compone de una serie de datos representados por puntos, unidos por segmentos lineales. se suele utilizar con variables cuantitativas, para ver su comportamiento en el transcurso del tiempo.





01.-Estadística Descriptiva. Gráficos



Diagrama: tipo de representación gráfica que sirve para representar un conjunto de datos.

A) Diagrama de barras: representar datos de variables cualitativas



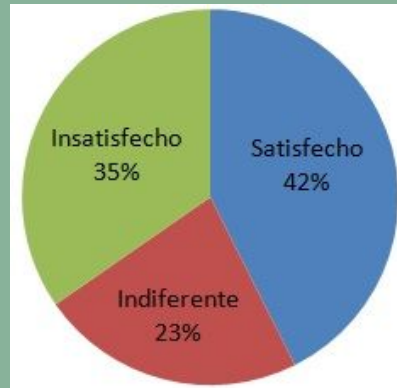


01.-Estadística Descriptiva. Gráficos



Diagrama: tipo de representación gráfica que sirve para representar un conjunto de datos.

Diagrama circular: proporción de elementos



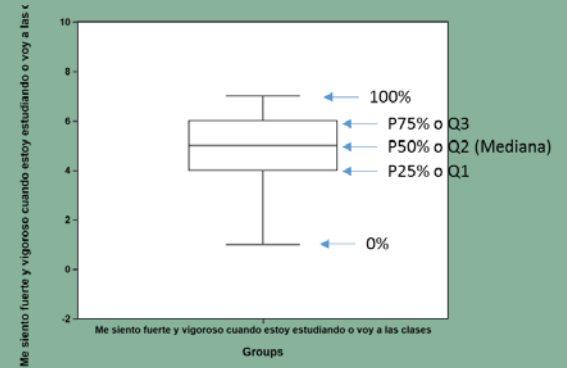
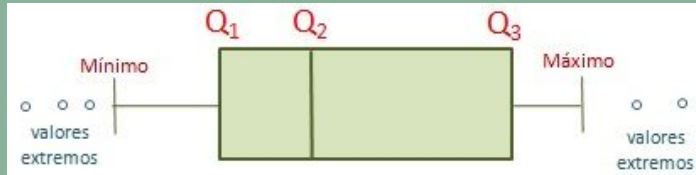


01.-Estadística Descriptiva. Gráficos



Diagrama: tipo de representación gráfica que sirve para representar un conjunto de datos.

Diagrama de caja: distribución en cuartiles



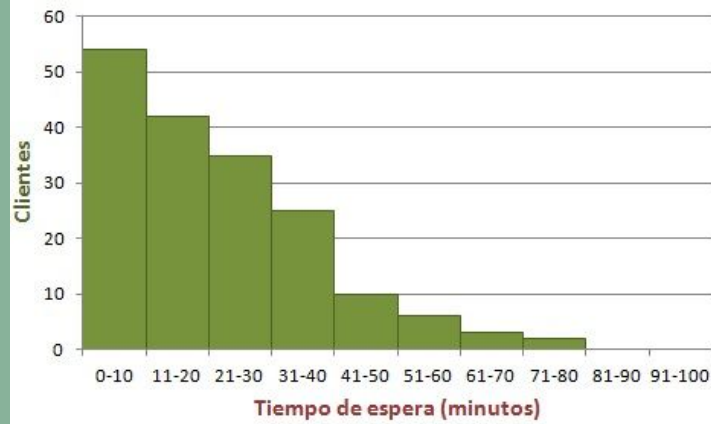


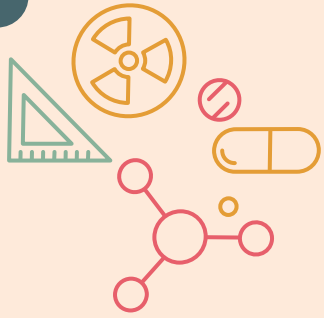
01.-Estadística Descriptiva. Gráficos



Diagrama: tipo de representación gráfica que sirve para representar un conjunto de datos.

Histograma: datos agrupados mediante intervalos





Índice de contenidos



01 Análisis descriptivo

02 Contraste de Hipótesis.

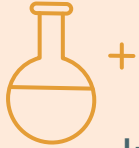
Capítulo 8 & 13; Patoja, A. (2009). Manual básico para la realización de tesinas, tesis y trabajos de investigación.

EOS Gabinete de Orientación Psicológica





02.-Contraste de Hipótesis. Introducción



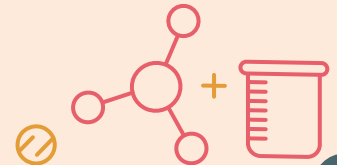
Ilustremos con un ejemplo

Imagina que te encuentras mal (escalofríos, dolor de cabeza, malestar general). Claros indicios que tienes fiebre, ¿verdad? Entonces piensas. Voy a comprobarlo. Te pones el termómetro y efectivamente marca 38°C. Ya sabes lo que toca. Antitérmico y sofá.

Esto es precisamente **un contraste de hipótesis!**

Tu **tienes una intuición** y quieres investigar si esta intuición es cierta.

- Tu **hipótesis de investigación** es que “tengo fiebre”. También llamada **hipótesis del investigador o alternativa (H_1)** Es la que quieres corroborar
- Por el contrario tienes **la hipótesis nula (H_0)**. ¿Por que nula?. Es la contraria. “NO tengo fiebre”



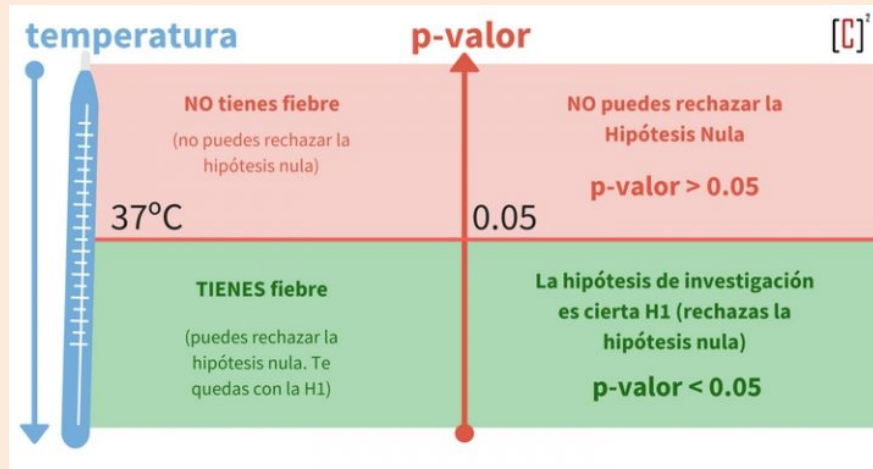


02.-Contraste de Hipótesis. Introducción



Bien pues cuando te pones el termómetro estás haciendo un **test estadístico**. Además utilizas un criterio para decidir si la hipótesis de investigación “tengo fiebre” es cierta.

- Si el valor es más alto que 37°C tienes fiebre.
- Por lo contrario no tienes fiebre





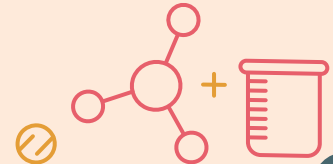
02.-Contraste de Hipótesis. Introducción

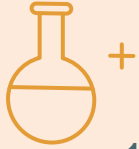


Qué es el p-valor: Un número que te va a dar el test estadístico para ver si te puedes quedar con la hipótesis de investigación o no. Pero necesitas un criterio. En el caso de la temperatura tienes el **criterio de 37°C**. En estadística normalmente se utiliza **el 0.05 (o el 5%)**

Sigue esta receta:

- **Si el p-valor es más grande que 0.05 NO puedes rechazar la hipótesis nula** (“NO tengo fiebre”)
- **Si el p-valor es más pequeño que 0.05 rechazas la hipótesis nula** (“NO tengo fiebre”) y **te quedas con la hipótesis de investigación** (“tengo fiebre”)





02.-Contraste de Hipótesis.

Pasos en el contraste de hipótesis



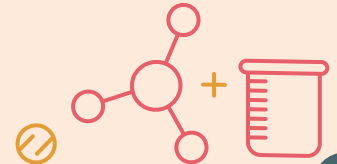
1. Formular las hipótesis de investigación

- H_1 o hipótesis del investigador (la que quieres justificar que es cierta con datos reales)
- H_0 o hipótesis nula (contraria que la H_1)

De manera aplicada:

Las notas de los alumnos de la clase A serán diferentes a la de la clase B

- H_1 : Nota Grupo A \neq Nota Grupo B (el efecto que espero)
- H_0 : Nota Grupo A = Nota Grupo B (el efecto contrario al esperado)





02.-Contraste de Hipótesis.

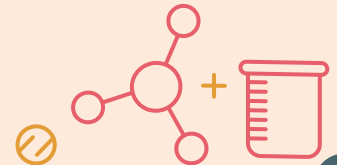
Pasos en el contraste de hipótesis



1. Formular las hipótesis de investigación
2. Decidir el criterio de significación (α):

El nivel de significación, (alfa o α) es la probabilidad de rechazar la hipótesis nula cuando es verdadera. Por ejemplo un nivel de significación de 0.05 indica un riesgo del 5% de concluir que existe una diferencia cuando no hay una diferencia real.

El criterio más típico es 0.05 (5%), así que para nuestro ejemplo usaremos este.





02.-Contraste de Hipótesis.



Pasos en el contraste de hipótesis

1. Formular las hipótesis de investigación
2. Decidir el criterio de significación (α): El criterio más típico es 0.05 (5%)
3. Elección de la prueba estadística (en base a nuestro diseño y nuestros datos)
 - ¿Qué test estadístico es el que vas a aplicar?

Dependerá del tipo de diseño y de los tipos de variables dependiente

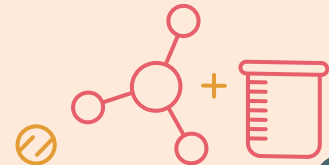
Contraste de medias

		Tipo de datos			
		Núméricos (gaussiana)	Ordinal o numérica (no gaussiana)	Núméricos (outliers)	Nominal binaria (2 resultados posibles)
Comparación	Comparar 2 grupos independientes	Prueba t para 2 muestras independientes	Prueba de Mann-Whitney	Prueba de Yuen para muestras independientes	Prueba de Fisher o Chi-cuadrado (para muestras grandes)
	Comparar 2 grupos relacionados	Prueba t para 2 muestras relacionadas	Prueba de Wilcoxon para muestras relacionadas	Prueba de Yuen para muestras relacionadas	Prueba de McNemar
	Comparar 3 o más grupos independientes	ANOVA de 1-vía para muestras independientes	Prueba de Kruskal-Wallis	ANOVA robusto de 1-vía para muestras independientes	Prueba Chi-cuadrado
	Comparar 3 o más grupos relacionados	ANOVA de 1-vía para muestras relacionadas	Prueba de Friedman	ANOVA robusto de 1-vía para muestras relacionadas	Prueba Q de Cochran
Asociar 2 variables		Correlación de Pearson	Correlación de Spearman o Kendall	Correlación robusta	Coefficiente V de Cramer

Católicas



Correlación





02.-Contraste de Hipótesis.

Pasos en el contraste de hipótesis



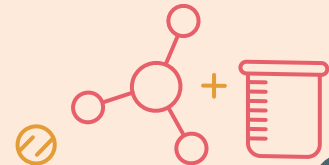
1. Formular las hipótesis de investigación
2. Decidir el criterio de significación (α): El criterio más típico es 0.05 (5%)
3. Elección de la prueba estadística (en base a nuestro diseño y nuestros datos)
4. Calcular el p-valor: Se calcula el valor de la prueba estadística que has decidido.

El resultado de nuestro estadístico $t = 3,305$ con un p -valor de 0.001

TEST RESULTS

t-test

Outcome	Group	Predictor	Descriptive statistics			Test statistics			Effect size		
			Group	n	Mean	Std. Dev.	df	t	p	Cohen's d	Hedge's g
Nota	Grupo		A	51	5.98	0.95	134.71	3.305	0.001	0.65	0.65
			B	87	5.30	1.47					





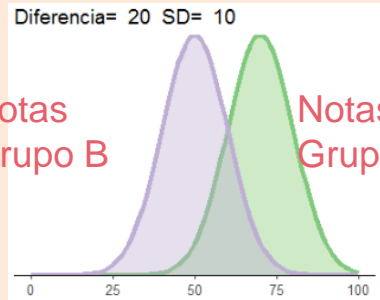
02.-Contraste de Hipótesis.

Pasos en el contraste de hipótesis

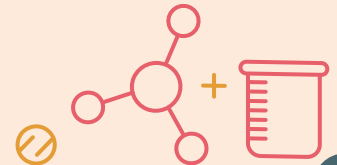


1. Formular las hipótesis de investigación
2. Decidir el criterio de significación (α): El criterio más típico es 0.05 (5%)
3. Elección de la prueba estadística (en base a nuestro diseño y nuestros datos)
4. Calcular el p-valor: Se calcula el valor de la prueba estadística que has decidido.
5. Conclusión | decisión:
 - $p < 0.05$ Rechazas la hipótesis nula y te quedas con la hipótesis del investigador..
 - $p > 0.05$ NO puedes rechazar la hipótesis nula y NO puedes demostrar que la hipótesis del investigador sea cierta. (NUNCA decir que la H_0 es cierta, sino que no la podemos rechazar)

$t = 3,305$, con un p -valor de 0.001



¿Qué decidimos?





Universidad de Jaén

Tema 5: Análisis Estadístico

Grado en Educación Infantil
Asignatura: Metodología de la Investigación en
Educación Infantil
Profesor: Samuel Parra León